1-2009

The Completeness of the Pragmatic Solution to Moore' Paradox: A Reply to Chan

John N. WILLIAMS
Singapore Management University, johnwilliams@smu.edu.sg

SMU SOCIAL SCIENCES & HUMANITIES
WORKING PAPER SERIES

SMU
SINGAPORE MANAGEMENT UNIVERSITY

The Completeness of the Pragmatic Solution to Moore' Paradox: A Reply to Chan

John Williams

July 2009

Paper No. 02-2009

The Completeness of the Pragmatic Solution to Moore's Paradox: A Reply to Chan

1. Introduction

In a recent paper in this journal, Timothy Chan (2008) argues that 'pragmatic' solutions to Moore's paradox are incomplete. Moore's paradox is the challenge of explaining why beliefs expressed by sentences of the form '*p* & I do not believe that *p*' ('Moorean sentences') are absurd, even though they (i) are contingent and often true, and (ii) express contents that are unproblematic when presented in the third person.[1] The 'pragmatic' solution turns upon the fact that believing a Moorean sentence falsifies it. If I believe that (*p* & I do not believe that *p*), then since believing a conjunction entails believing its conjuncts, I believe that *p*, which falsifies the content of my belief that (*p* & I do not believe that *p*).[2] Let us follow Chan in saying that such a belief is 'self-refuting'. By contrast, if the belief is transposed into the third person (or into any non-first-person) then it is not self-refuting. If I believe that (*p* & he does not believe that *p*) then since believing a conjunction involves believing its conjuncts, I believe that *p*, but this does not falsify the content of my belief that (*p* & he does not believe that *p*), because the fact that I believe that *p* is consistent with the fact that he does not believe that *p*.

Chan points out that the fact that holding a Moorean belief falsifies it is not enough to make the believer irrational (and in this sense, absurd). What is needed is that the believer can be expected to know that holding it falsifies it. Chan argues that more precise formulations of this ground of Moorean absurdity are either too weak in making

[1] Moore observes that to say, 'I went to the pictures last Tuesday but I don't believe that I did' would be 'absurd' (1942, p. 543).
[2] See Williams 1994; 1996; 1998; 2006.

beliefs that are not Moorean count as Moorean or too strong in making beliefs that are Moorean count as non-Moorean. He concludes that the only formulation that is neither too weak nor too strong has to be couched in first-person terms, making an essential use of 'I'. He thinks that this means that the pragmatic solution *presupposes* the asymmetry between first- and non-first-person expressions of the content of the belief. The pragmatic solution therefore cannot explain the asymmetry. Since the asymmetry is a crucial feature of the paradox, the pragmatic solution is incomplete, because it fails to account, in a non-circular way, for (ii)—why Moorean beliefs cease to be absurd when their contents are presented in the third person.

I will show that Chan's argument fails in two ways. Firstly, it is unclear why there really should be any circularity. Secondly, there is a formulation of the ground of Moorean absurdity—one that Chan himself considers and rejects as being too strong—that is neither too weak nor too strong and that does not presuppose the asymmetry between first- and non-first-person expressions of the content of the Moorean belief. Thus the pragmatic solution to Moore's paradox is still in good shape.

2. Chan's Argument

Let us now trace the steps in Chan's argument. He starts by comparing Moorean belief with beliefs in necessary falsehoods. He observes that the fact that one believes a necessary falsehood is not sufficient for one to be irrational. For example, mathematicians before Gödel were not irrational in believing that arithmetic is decidable, because they could not have been expected to know that it is necessarily false that

arithmetic is decidable. Analogously, the fact that one's belief is self-refuting is not a sufficient condition for one to be irrational, because one might not be reasonably expected to know that it is self-refuting. For example, one is not irrational in believing that

> The atheism of my mother's nieceless brother's only nephew angers God

if one reasonably fails to see that it entails

> God exists but I don't believe that God exists.[3]

In order to take this point into account, the pragmatic solution must define the *Ground* of the *Absurdity* of *Moorean* beliefs as follows:

> (GAM) If an agent could be expected to know that necessarily (if q is believed, then $\sim q$), then it would be absurd to for him to believe that q.

Of course (GAM) is open to alternative interpretations. One of these, when formalized (hence the '*f*') is:

> (GAM*f*.1) $\forall x(\mathrm{K}x\Box(\mathrm{B}xq \rightarrow \sim q)) \rightarrow \mathrm{AB}xq)$
>
> It would be absurd for an agent to believe that q, if he could be expected to know that necessarily he could not truly believe that q.[4]

where 'x' ranges over possible agents, '$\mathrm{K}xq$' abbreviates 'x could be expected to know that q', meaning it would be very unreasonable for x not to know that q; and '$\mathrm{AB}xq$' abbreviates 'it would be absurd for x to believe that q'.

Chan points out that this too weak. Gordon Brown could be expected to know that Gordon Brown cannot truly believe that (Gordon Brown will lose the election but he doesn't believe it), but if Gordon Brown does not know that he *is* Gordon Brown

[3] This example is given by Roy Sorensen (1988, p.28).

[4] I have a qualm about Chan's translation, one that applies to the rest of his translations as well. This is that that the fact that someone cannot truly believe that q, is entailed by, but does not entail, the fact that this persons' holding the belief thereby falsifies it. I am not sure if this affects Chan's argument, but surely the second fact is more central to the absurdity.

(perhaps because he is suffering from amnesia) there is no absurdity in him believing that (Gordon Brown will lose the election but he doesn't believe it). In order to single out those propositions that are obviously necessarily false-if-believed-by-*x* independently of *x*'s knowledge of who he is, we might try:

(GAM*f*.2) $\forall x(\mathrm{K}x\Box(\forall y(\mathrm{B}yq \rightarrow \sim q)) \rightarrow \mathrm{AB}xq)$

> It would be absurd for an agent to believe that *q*, if he could be expected to know that no one could truly believe that *q*.

Chan claims that this is too strong. To see this, suppose that '*q*' represents '*p* and I do not believe that *p*' and let us instantiate '*x*' as *a*, to give:

> $\mathrm{K}a\Box\forall y(\mathrm{B}y($*p* and I do not believe that *p*$) \rightarrow \sim ($*p* and I do not believe that *p*$))$
> $\rightarrow \mathrm{AB}a($*p* and I do not believe that *p*$)$

Chan provides no English translation. It would be:

> If *a* could be expected to know that no one could truly believe that (*p* and I do not believe that *p*) then *a* is absurd in believing that (*p* and I do not believe that *p*).

Chan observes that the antecedent of this conditional must be false.[5] This is because someone other than me, say Bill, might truly believe that (*p* and *I* do not believe that *p*), if his belief is *about me* and not about himself. Therefore it is false that no one could truly believe that (*p* and I do not believe that *p*). Since nobody can be expected to know what is false, it is false that *a* could be expected to know that no one could truly believe that (*p* and I do not believe that *p*).

Chan draws the moral that we need to use the context-dependence of the semantic value of indexicals such as 'I'; what (GAM*f*.2) really means to say is that *whoever* utters a (token of a) Moorean sentence, and thus *whatever* proposition it expresses, that

[5] An exception is when it is necessarily false that *p*.

proposition could not be truly believed by the utterer. This means that (GAM) should not be formulated in terms of the relation

> Bxq iff$_{\text{def}}$ x believes that q.

which holds between the agent x and the *proposition* that q but rather in terms of a relation between an agent and a *sentence*:

> B$_s x$‘q’ iff$_{\text{def}}$ x holds a belief which x would express by (assertively) uttering the sentence ‘q’ in x’s current context, if he wished to express it.

This is more fine-grained than

> B$_{ps} x$‘q’ iff$_{\text{def}}$ x believes *that* the proposition expressed by ‘q’ in x’s mouth in x’s current context is true.

which holds between an agent and the proposition expressed by a sentence. Accordingly we may formulate (GAM) as

> (GAMf.3) $\forall x(\mathrm{K}x\Box(\forall y(\mathrm{B}_{ps}y\text{‘}q\text{’} \rightarrow \text{‘}q\text{’ is false})) \rightarrow \mathrm{AB}_s x\text{‘}q\text{’})$
>
> It would be absurd for an agent to hold a belief which he would express by assertively uttering ‘q’, if he could be expected to know that ‘q’ could not express a true belief in anybody’s mouth in any context.

Chan offers an example that is supposed to show that this is too weak: Suppose that I am having a debate with John Smith on MSN Messenger, trying to convince him that the Earth is well over 5,000 years old. The screen is divided into two halves, labelled ‘John Smith and ‘NN’ [author’s initials]. After a while I notice that whatever I type appears on his screen. Thinking that Smith is mimicking my words, I try to catch him out by typing

> (T) The person actually typing these very words now here on the upper half of my screen does not believe that the Earth is well over 5,000 years old, but of course it is.

These words then appear on John Smith’s screen. But unknown to me, I am the person typing the words on his screen, because the system is malfunctioning. There is no

absurdity in my belief of (T). But Chan argues that (GAM*f*.3) predicts that my belief in (T) is absurd, and so is too weak. He says:

> Now I knew that whoever is referred to by 'the person actually typing …' be he John or myself or anyone else, the token of [(T)] that I was looking at on the top half of my screen as I typed must be false if its producer believed the proposition it expressed in his current context. I did believe the proposition it expressed in my context, and it was (unknown to me) I who was producing it, so it could not have been true.

My belief is not absurd because it is not irrational for me to fail to know, when I produce the sentence token on the upper half of my screen, that its producer was *me*. Thus using '*i*' to abbreviate the word 'I', and employing the stipulative definition

> $K_s x$'q' iff$_{\text{def}}$ x knows what x would express by (assertively) uttering the sentence 'q' in x's current context, if he wished to express it.

we need

(GAM*f*.4) $\forall x(K_s x$"$\Box(B_{ps} i$'q'$\rightarrow$ 'q' is false)" $\rightarrow AB_s x$'q'$)$

> It would be absurd for an agent to hold a belief which he would express by assertively uttering 'q', if he could be expected to know, "I could not truly believe the proposition expressed by 'q' in my mouth in my current context."

Chan observes that this is equivalent to

(GAM*f*.4*) $\forall x(Kx\Box(B_{ps} x^*$'$q$' $\rightarrow$ 'q' is false) $\rightarrow AB_s x$'q'$)$

> It would be absurd for an agent to hold a belief which he would express by assertively uttering 'q', if he could be expected to know that he* could not truly believe the proposition expressed by 'q' in his* mouth in his* current context

where '$x^*\varphi$' abbreviates 'x believes that he himself φ'.[6]

---

[6] This is an extension of Castañeda's (1966; 1968) quasi-indicator notation, according to which 'x believes that $x^*\Phi$' is read as 'x believes that he himself Φ'. 'x^* Φ' falls within the scope of a propositional attitude

This is neither too weak nor too strong. But (GAM*f*.4) requires that *x* would express his knowledge by uttering, "… *I* believe the proposition expressed by '*q*' …". Likewise (GAM*f*.4*) requires that *x* conceives of himself in first-personal terms. The problem now is supposed to be that these formulations exhibit first-/third-person asymmetry. If we replace '*i*' in (GAM.*f*4) with any other co-referring, non-first-person expression, or in other words delete '*' from (GAM*f*.4*), we get

(GAM*f*.5) $\forall x(\mathrm{K}x\Box(\mathrm{B}_{ps}x\text{‘}q\text{’} \rightarrow \text{‘}q\text{’ is false}) \rightarrow \mathrm{AB}_{s}x\text{‘}q\text{’})$

> It would be absurd for an agent to hold a belief which he would express by assertively uttering '*q*', if he could be expected to know that he could not truly believe the proposition expressed by '*q*' in his mouth in his current context.

Chan points out that the same example that shows (GAM.*f*1) to be too weak, namely that of the amnesiac Gordon Brown, also shows this to be too weak as well: Gordon Brown could be expected to know that Gordon Brown could not truly believe the proposition expressed by 'Gordon Brown will lose the election but he doesn't believe it' in his mouth in his current context. But if Gordon Brown cannot be expected to know that he *is* Gordon Brown then there is no absurdity in his holding a belief which he would express by assertively uttering 'Gordon Brown will lose the election but he doesn't believe it'.

Strengthening the condition defined by (GAM*f*.5) to get:

---

attributed to *x*, as in '*x* believes that *x** is walking with a stoop' or '*x* fears that *x** is unattractive to women'. Not all uses of the reflexive pronouns 'himself' and 'myself', can be parsed as '*x**', for example, 'He cut himself accidentally' and 'I disqualified myself on purpose'. An alternative translation of (GAM*f*.4*) into English that is perhaps more perspicuous is:

> It would be absurd for an agent to hold a belief which he would express by assertively uttering '*q*', if he could be expected to know that *he himself* could not truly believe the proposition expressed by '*q*' in *his very own* mouth in *his very own* current context.

(GAM*f*.6) $\forall x(\mathrm{K}x\Box(\forall y(\mathrm{B}_{ps}y\text{‘}q\text{’} \rightarrow \text{‘}q\text{’ is false})) \rightarrow \mathrm{AB}_s x\text{‘}q\text{’})$

> It would be absurd for an agent to hold a belief which he would express by assertively uttering ‘*q*’, if he could be expected to know that ‘*q*’ could not express a true belief in anybody’s mouth in any context

is no help, because (GAM*f*.6) is identical to (GAM*f*.3), which Chan thinks is too weak. So the ‘*’ in (GAM*f*.4*) is essential for its truth. Therefore the use of the first person in the formulation of (GAM), a principle needed by the pragmatic solution, is not reducible to the third person. Chan concludes that the pragmatic solution cannot explain why Moorean beliefs cease to be absurd when their contents are presented in the third person.

3. The Failure of Chan’s Argument

Chan’s argument faces two objections. The first is that is that it is unclear why the fact that the pragmatic explanation of the absurdity of Moorean belief must be framed in terms of the first person means that it cannot explain, in a non-circular way, why Moorean beliefs cease to be absurd when their contents are presented in the third person. Circularity arises only if the pragmatic explanation of the absurdity presupposes that the absurdity *disappears for* third-person beliefs. But this disappearance does not seem to be presupposed by the fact that the absurdity *cannot be explained in terms of* the third-person. To put the point another way, an necessary condition of circularity in explanation is that the *explanandum* entails the *explanans*. But the fact that Moorean beliefs are only absurd in the first-person does not appear to entail statements such

(GAM*f*.4) $\forall x(K_s x \text{“}\Box(B_{ps} i \text{‘}q\text{’} \rightarrow \text{‘}q\text{’ is false)”} \rightarrow AB_s x \text{‘}q\text{’})$

> It would be absurd for an agent to hold a belief which he would express by assertively uttering ‘*q*’, if he could be expected to know, “I could not truly believe the proposition expressed by ‘*q*’ in my mouth in my current context.”

or equivalently

(GAM*f*.4*) $\forall x(Kx\Box(B_{ps} x^* \text{‘}q\text{’} \rightarrow \text{‘}q\text{’ is false}) \rightarrow AB_s x \text{‘}q\text{’})$

> It would be absurd for an agent to hold a belief which he would express by assertively uttering ‘*q*’, if he could be expected to know that he* could not truly believe the proposition expressed by ‘*q*’ in his* mouth in his* current context.

The result is that it remains unclear why these are not satisfactory explanations of the absurdity of Moorean beliefs.

The second problem with Chan’s argument concerns his discussion of

(GAM*f*.3) $\forall x(Kx\Box(\forall y(B_{ps} y \text{‘}q\text{’} \rightarrow \text{‘}q\text{’ is false})) \rightarrow AB_s x \text{‘}q\text{’})$

> It would be absurd for an agent to hold a belief which he would express by assertively uttering ‘*q*’, if he could be expected to know that ‘*q*’ could not express a true belief in anybody’s mouth in any context.

Call t_1 the time at which I type

> (T) The person actually typing these very words now here on the upper half of my screen does not believe that the Earth is well over 5,000 years old, but of course it is.

and the time at which I read (T) on John Smith’s screen, t_2. Clearly I don’t believe the proposition that (T) expresses at t_1. After all, I’m not using (T) to assert anything, only

typing (T) in order that John Smith will type (T) in return at t_2. If I did believe (T) at t_1 then I'd know that I was the person referred to by 'the person actually typing …' In that case, if I did believe the proposition (T) expresses, my belief would be absurd. However, there is no absurdity if I believe the proposition (T) expresses at t_2. But this does not falsify (GAM*f*.3) as Chan states it in English. This is because it is false that (T) could not express a true belief in *anybody's* mouth in *any* context. In fact at t_2, it might express a true belief in *my* mouth in *my* context. This is the whole point of my attempt to catch John Smith out, because were John Smith to produce (T) assertively then he would be guilty of Moorean absurdity, and all the more so if he actually believed what (T) expresses. Since it is false that (T) could not express a true belief in anybody's mouth in any context, I could not be expected to know that (T) could not express a true belief in anybody's mouth in any context, because I cannot be expected to know what is false. This means that the antecedent of (GAM*f*.3) is false, so (GAM*f*.3) does not predict absurdity where there is none. Another way to put this point is that I did believe the proposition (T) expressed at t_2 and it was (unknown to me) I who was producing it, but it does not follow that it could not have been true, because it need not have been me, but John Smith that was producing it.

Chan might respond by conceding that his English translation of (GAM*f*.3) is not accurate, and that a better translation is

> If you could be expected to know that anyone who holds a belief expressed by '*q*' in *his* mouth in *his* current context, thereby falsifies that belief, then you would be absurd to hold the belief that you would express by (assertively) uttering '*q*'in your current context.

But this move falls to the much the same objection. It is false that anyone who holds a belief expressed by '*q*' in *his* mouth in *his* current context thereby falsifies that belief. Suppose that at t_2, I hold the belief expressed by

> 'The person actually typing these very words now does not believe that the Earth is well over 5,000 years old although the Earth is well over 5,000 years old'

in my mouth in my current context. Suppose also that John Smith is silly enough to have fallen into my trap and it is he who is actually doing the typing. Then my belief might be true. Since it is false that anyone who holds a belief expressed by (T) in *his* mouth in *his* current context, thereby falsifies that belief, I could not be expected to know that anyone who holds a belief expressed by (T) in *his* mouth in *his* current context, thereby falsifies that belief, because I cannot be expected to know what is false. So again the antecedent of (GAM*f*.3) is false and so again, does not predict absurdity where there is none.

We may conclude that (GAM*f*.3) is neither too weak nor too strong an explanation of the absurdity of Moorean beliefs. Moreover it presupposes no asymmetry between first- and non-first-person expressions of the content of the Moorean beliefs. Furthermore (GAM*f*.4) or equivalently, (GAM*f*.4*), are also neither too weak nor too strong an explanation, and although they are framed in terms of the first-person, there is no reason to think that this introduces circularity into their explanation of why Moorean beliefs cease to be absurd when their contents are presented in the third person.

The upshot is that the pragmatic solution to Moore's paradox is still in good shape.[7]

[7] I thank Claudio de Almeida for useful discussion

References

Castañeda, H. N. 1966. He*, a Study in the Logic of Self-Consciousness. *Ratio* 6: 130–157.

_____ 1968. On the Logic of Attribution of Self-Knowledge to Others. *Journal of Philosophy* 65: 439–56.

Chan, T. 2008. Moore's Paradox Is Not Just Another Pragmatic Paradox, *Synthese* (online first, October).

Moore, G. E. 1942. A Reply to My Critics in P. Schlipp (ed.) *The Philosophy of G. E. Moore*, Tudor, Evanston, 535-677.

Sorensen, R. A. 1988. *Blindspots*. (Oxford: Clarendon)

Williams, J. N. 1994. Moorean Absurdity and the Intentional 'Structure' of Assertion, *Analysis* 54.3**,** 60-66.

_____1996. Moorean Absurdities and the Nature of Assertion, *Australasian Journal of Philosophy* 74.1, 115-130.

_____1998. Wittgensteinian Accounts of Moorean Absurdity, *Philosophical Studies* 92.3, 283-306.

_____2006. Wittgenstein, Moorean Absurdity and its Disappearance from Speech, *Synthese* 149.1, 225-254.